



Hábitos de búsqueda de información científica e influencia en la visibilidad de las revistas

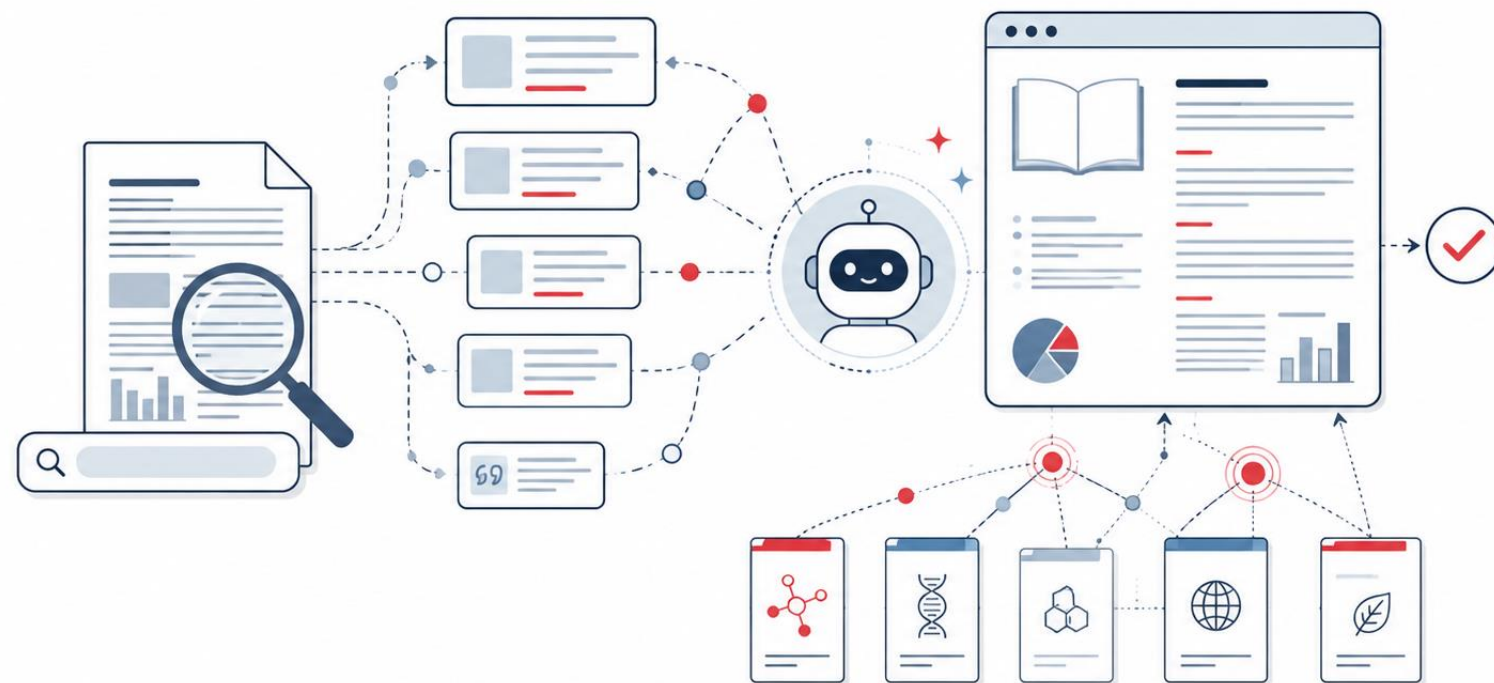


25 de septiembre de 2026

Enrique Orduña-Malea 
Universitat Politècnica de València



Analizar cómo la incorporación de asistentes de investigación está transformando los hábitos de búsqueda de información científica, con **potenciales efectos sobre la selección, visibilidad y citación** de las publicaciones y de las revistas científicas que las difunden.



A

DELIMITAR los principales **asistentes de investigación**, diferenciando entre recuperación tradicional y *deep research*.

B

CARACTERIZAR los **hábitos actuales de búsqueda de información científica** y los factores que impulsan su transformación, especialmente el uso creciente de herramientas de IA.

C

ANALIZAR *Ai2 Asta*, atendiendo a sus procesos de recuperación, selección, síntesis y citación, y a la **presencia de revistas** en sus resultados.

D

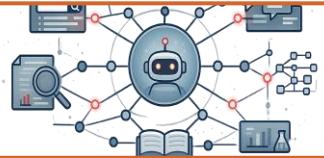
EXAMINAR sus **implicaciones para la comunicación científica**, especialmente sobre transparencia, reproducibilidad, sesgos y distribución de la visibilidad y el impacto entre publicaciones y revistas.



Un ecosistema de búsqueda en transformación



De la búsqueda al agente: *deep search* y *deep research*



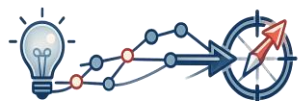
El nuevo ecosistema de asistentes de investigación



Cómo buscan hoy los investigadores



Ai2 Asta: síntesis de literatura y sistema de citas



Lecciones aprendidas, implicaciones y futuro

1. Un ecosistema de búsqueda en transformación

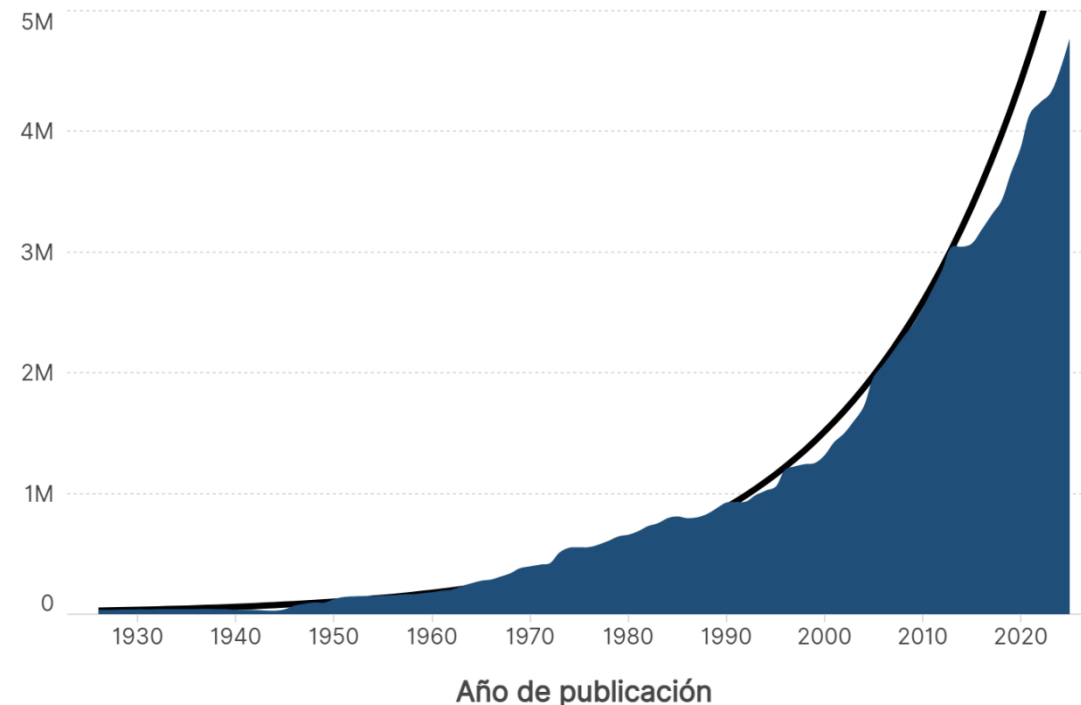


5

La búsqueda científica como necesidad creciente

- La **producción científica crece y crece**, aumentando el volumen y la dispersión de la literatura potencialmente relevante.
- Buscar información se ha convertido en una **actividad continua y transversal**, necesaria para actualizarse, revisar la literatura, preparar proyectos y desarrollar nuevas investigaciones.
- El **tiempo dedicado a la búsqueda ha aumentado desde 2020**, especialmente entre el personal de investigación en la primera etapa de su carrera.

Número de publicaciones



Dedicar **más tiempo no garantiza necesariamente encontrar mejor**: las formas actuales de buscar y presentar resultados también introducen fricciones y limitaciones.



Limitaciones de la búsqueda académica pre-IA

Sobrecarga de información:

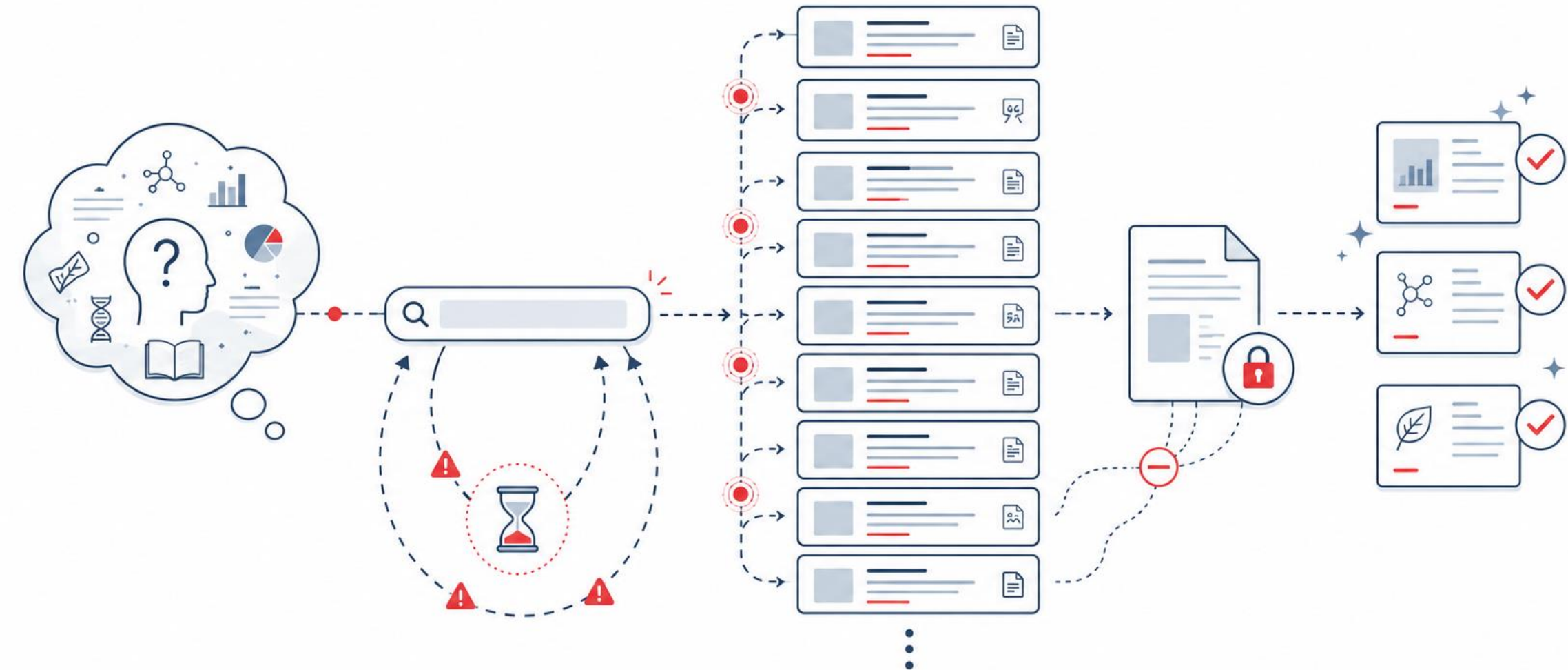
Existen demasiados resultados relevantes, lo que exige filtrar. Para ello se requiere de **tiempo, esfuerzo y criterio**.

Limitaciones en la forma de buscar:

Las búsquedas **dependen de la selección de términos**, lo que exige traducir una necesidad compleja en una consulta breve, en lenguaje natural pero restringida y con poca capacidad de ser realmente comprendida.

Limitaciones de las herramientas:

Los resultados ofrecidos **no siempre son precisos o contextualizados** y la relevancia depende de la **cobertura** de la herramienta y de los **algoritmos** utilizados para ordenar los resultados.



El modelo pre-IA mantiene numerosas fricciones y transfiere al investigador buena parte del trabajo de **reformulación**, **filtrado** y **evaluación**.

1. Un ecosistema de búsqueda en transformación



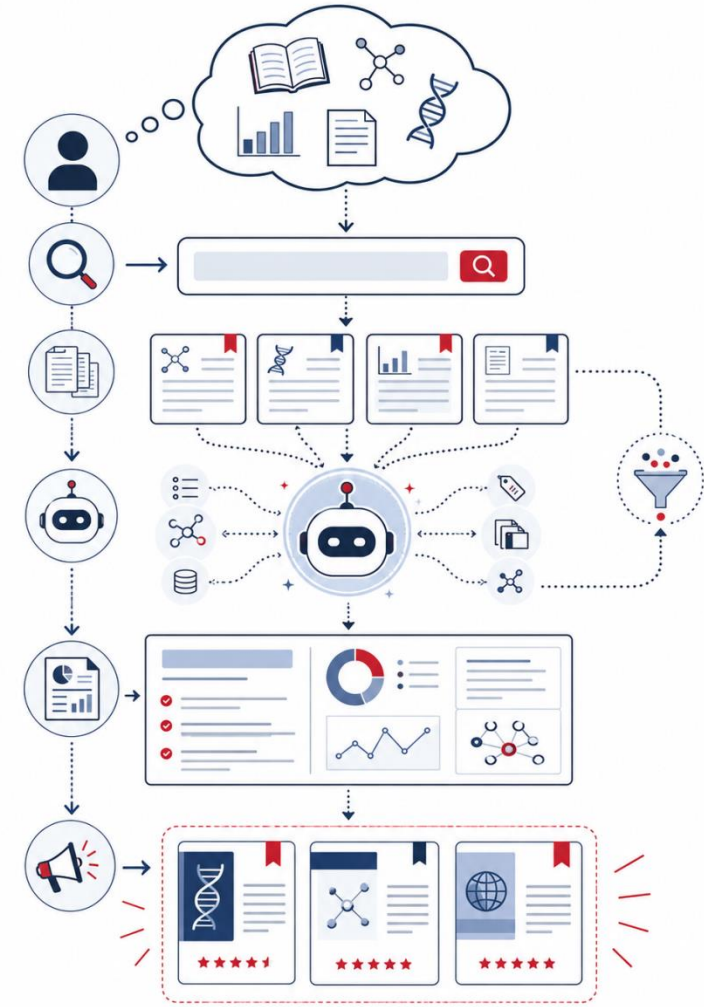
8

De la búsqueda tradicional a la búsqueda asistida por IA

Las herramientas basadas en IA introducen **nuevas formas de formular preguntas y explorar la literatura**.

Permiten pasar **de consultas breves por palabras clave a instrucciones** más complejas en lenguaje natural.

Estas herramientas no se limitan a recuperar documentos: algunas también **seleccionan, organizan, sintetizan y presentan** resultados.



El cambio no es sólo tecnológico, sino funcional: ya no hablamos únicamente de encontrar documentos, sino de **sistemas que comienzan a intervenir en la selección, organización e interpretación de la literatura científica**.

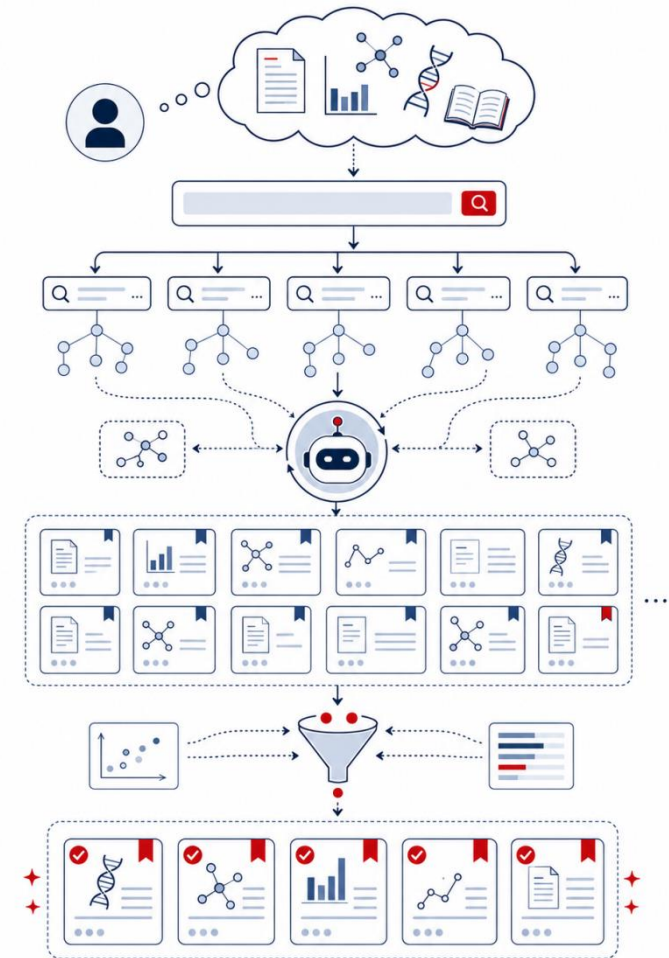
2. De la búsqueda al agente



9

Deep search: más allá de la búsqueda de una sola iteración

- El *deep search* amplía la búsqueda tradicional, combinando **lenguaje natural y búsqueda semántica (uso de *embeddings*)**
- No se limita a ejecutar una única consulta:** puede descomponer la necesidad informativa en subconsultas, explorar varias rutas y refinar progresivamente los resultados.
- Representa un paso hacia la búsqueda *agentic*, porque parte del trabajo de exploración, reformulación y selección **deja de recaer por completo en la persona investigadora.**



El *deep search* no cambia sólo cómo se consulta: **cambia cómo el sistema busca, explora y selecciona evidencias** antes de devolvérsela al usuario.



Deep research: de la recuperación a la generación de respuestas e informes

■ Síntesis de acciones

El *deep research* añade una capa de síntesis sobre la búsqueda, integrando recuperación documental **con la generación de respuestas fundamentadas**.

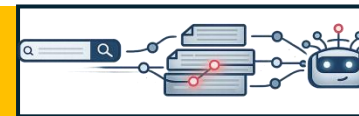
■ Informes fundamentados

Estos sistemas pueden **comparar estudios, resumir hallazgos u organizar evidencias**.

■ Efectos en el ecosistema científico

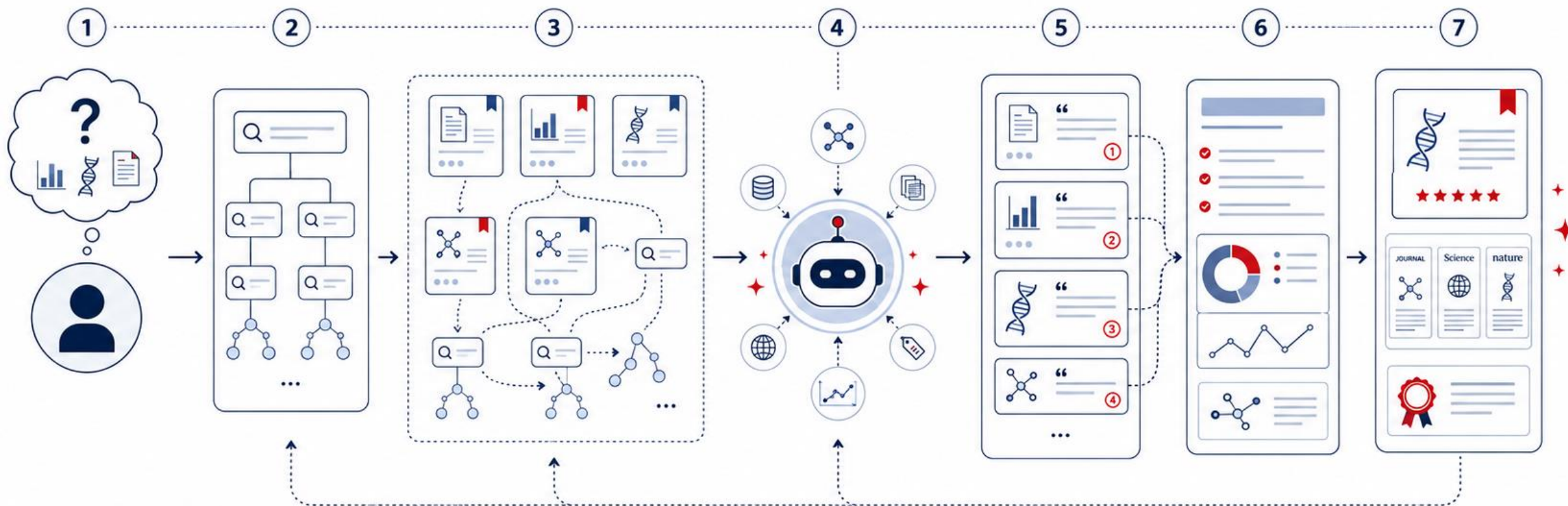
No solo facilitan el descubrimiento de literatura, sino que **pueden influir en qué publicaciones y qué revistas quedan incorporadas a la respuesta final**.

2. De la búsqueda al agente



11

Deep research: de la recuperación a la generación de respuestas e informes



El *deep research* desplaza el foco desde la mera recuperación de documentos hacia la **producción de una respuesta o informe sustentado en evidencia**.

3. El nuevo ecosistema de asistentes de investigación



12

Clasificando el nuevo ecosistema de búsqueda académica

Quick Search:

Búsqueda de una sola iteración; devuelve una lista de documentos relevantes.

Quick RAG:

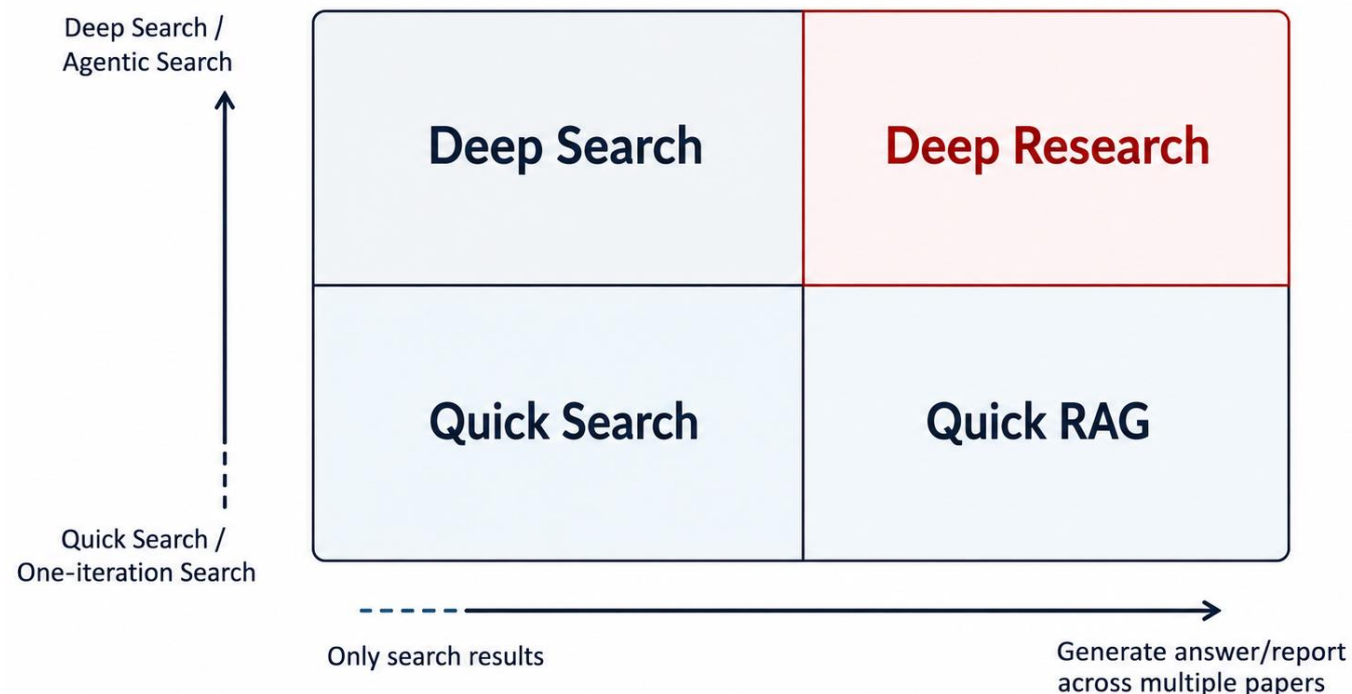
Una recuperación simple sustenta una respuesta breve generada con apoyo documental.

Deep Search:

Una búsqueda iterativa y *agentic* mejora la exploración, cobertura y priorización de resultados.

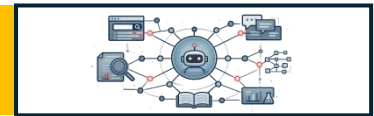
Deep Research:

Combina búsqueda profunda, selección de evidencias y generación de respuestas estructuradas.



Fuente: elaboración propia a partir de Aaron Tay (2025a, 2025b).

3. El nuevo ecosistema de asistentes de investigación



13

Una plataforma, distintas funciones de búsqueda y asistencia

Las fronteras entre cuadrantes son **permeables**: una misma plataforma puede integrar búsqueda rápida, RAG, *deep search* y *deep research* mediante modos diferentes.

La evolución del ecosistema es **acumulativa y modular**: las nuevas herramientas no siempre sustituyen a las anteriores, sino que añaden capas de búsqueda semántica, iteración, síntesis y generación de respuestas.

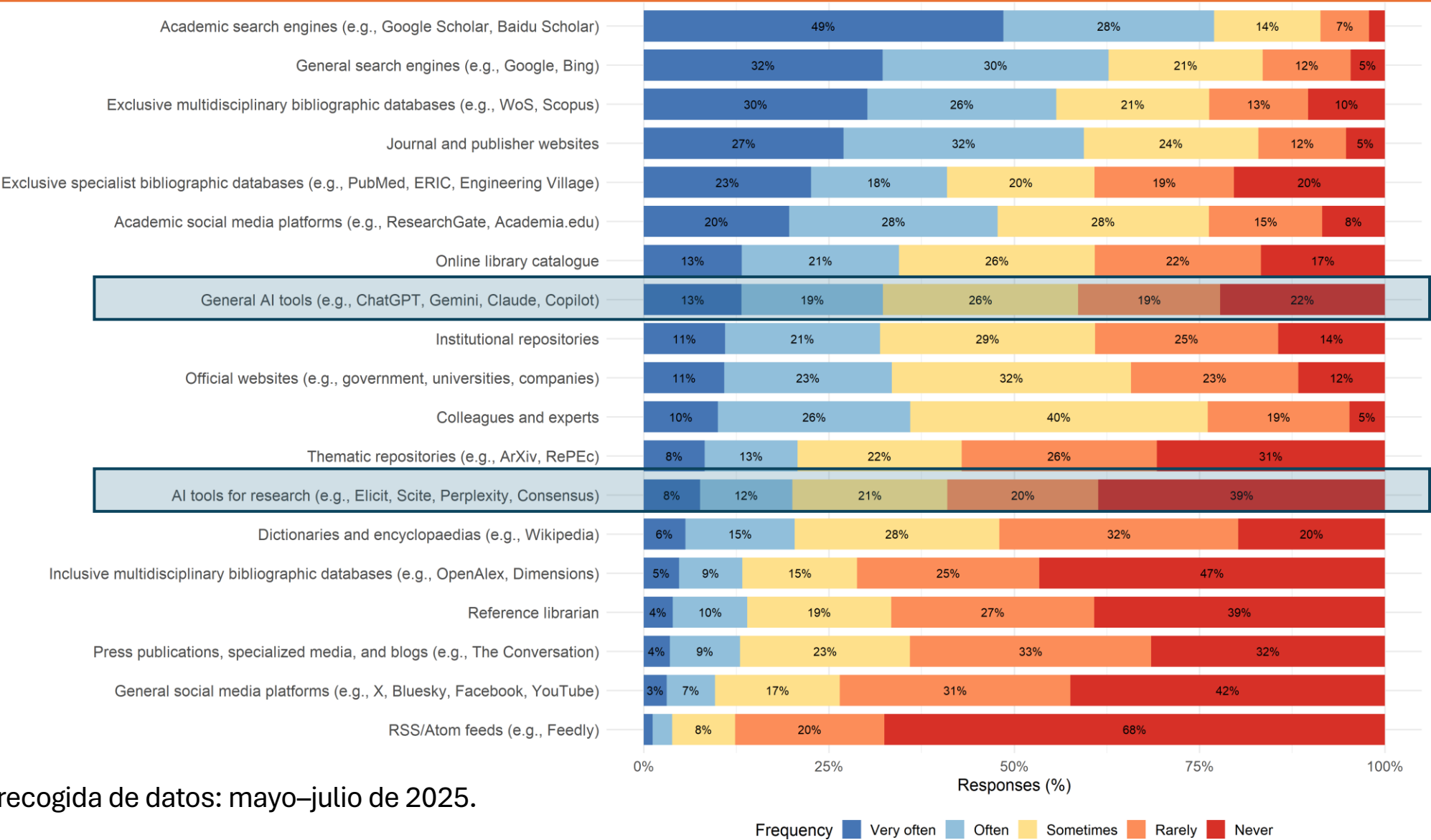
Cuanto mayor es la intervención del sistema, **mayor puede ser su efecto sobre la visibilidad científica**, porque ya no solo ordena documentos: también decide cuáles utiliza para construir y respaldar la respuesta final.

Más que clasificar plataformas completas, debemos identificar **qué función ejecuta cada herramienta y cuánto interviene entre la pregunta del investigador y la evidencia finalmente visible**.



4. Cómo buscan hoy los investigadores

¿Qué herramientas de búsqueda de información científica utilizas?



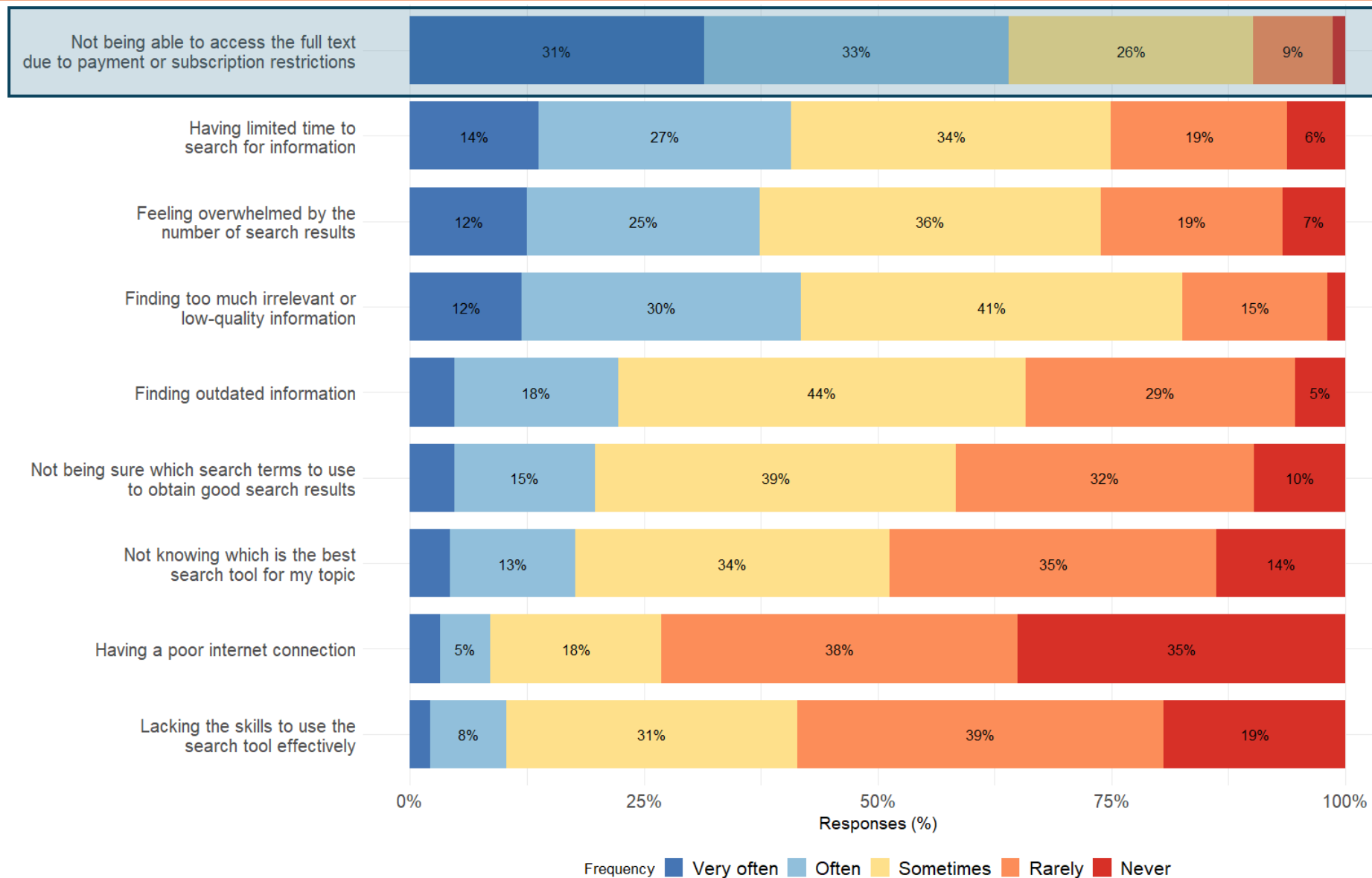
Periodo de recogida de datos: mayo–julio de 2025.
n=3.349

4. Cómo buscan hoy los investigadores



15

¿Cuáles son las principales limitaciones a la hora de buscar información científica?





A scholarly research assistant that combines literature understanding and data-driven discovery.

Asta uses 108M+ abstracts and 12M+ full-text papers to find, summarize, and analyze scientific evidence. A project from [Ai2](#).

Find papers

Generate a report

Analyze data

Ask a scientific question. Use "@" to mention a paper.

Thinking mode ▾

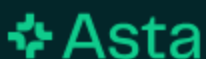


Show example queries

Explore research prototypes

<https://asta.allen.ai/>



 > Generate a report

Share  Sign Out 

U How are university rankings constructed based on bibliometric indicators, and what are the main uses, quality concerns, biases, and advantages or disadvantages of these rankings?

 Researched for 2m 44s [View steps](#) 

 Your report is ready!
[View Report](#) 

Are these answers helpful?    [Download PDF](#) 

 **Bibliometric University Rankings: Construction, Uses, and Quality Concerns**

▼ **Background and Emergence**

University rankings based on bibliometric indicators emerged in the mid-2000s, starting with the Shanghai Ranking in 2003, driven by global competition and the knowledge economy. These rankings primarily measure research performance and reputation using citation databases, though they face significant methodological criticisms from the bibliometric community.

University rankings emerged during the mid-2000s in the context of the emerging knowledge-based economy and increasing competition for world-class universities ([Leydesdorff et al., 2010](#)). The launch of the Shanghai Ranking in 2003 marked a pivotal moment, not only setting up the



Metodología



Fichero: <https://huggingface.co/datasets/allenai/asta-summary-citation-counts>



5. Ai2 Asta: síntesis de literatura y sistema de citas

Datos descriptivos generales del sistema de citación de Asta

Indicador	Valor
Fuentes normalizadas	80.667
Publicaciones citadas totales	2.357.300
Publicaciones citadas con fuente identificada	1.942.058
Apariciones en consultas	3.431.508
Citas totales	4.816.003

La literatura citada por *Asta* procede de un **universo muy amplio y diverso** de fuentes de publicación.
17.6% de las publicaciones no están vinculadas a una fuente de publicación



5. Ai2 Asta: síntesis de literatura y sistema de citas

Concentración de fuentes

Grupo	Nº de fuentes	Publicaciones citadas	Apariciones en consultas	Citas
Top 1 %	807	49,7 %	52,3 %	52,6 %
Top 10 %	8.067	80,8 %	82,4 %	82,8 %
Top 25 %	20.167	91,6 %	92,9 %	93,1 %
Top 50 %	40.334	97,1 %	97,9 %	98,0 %

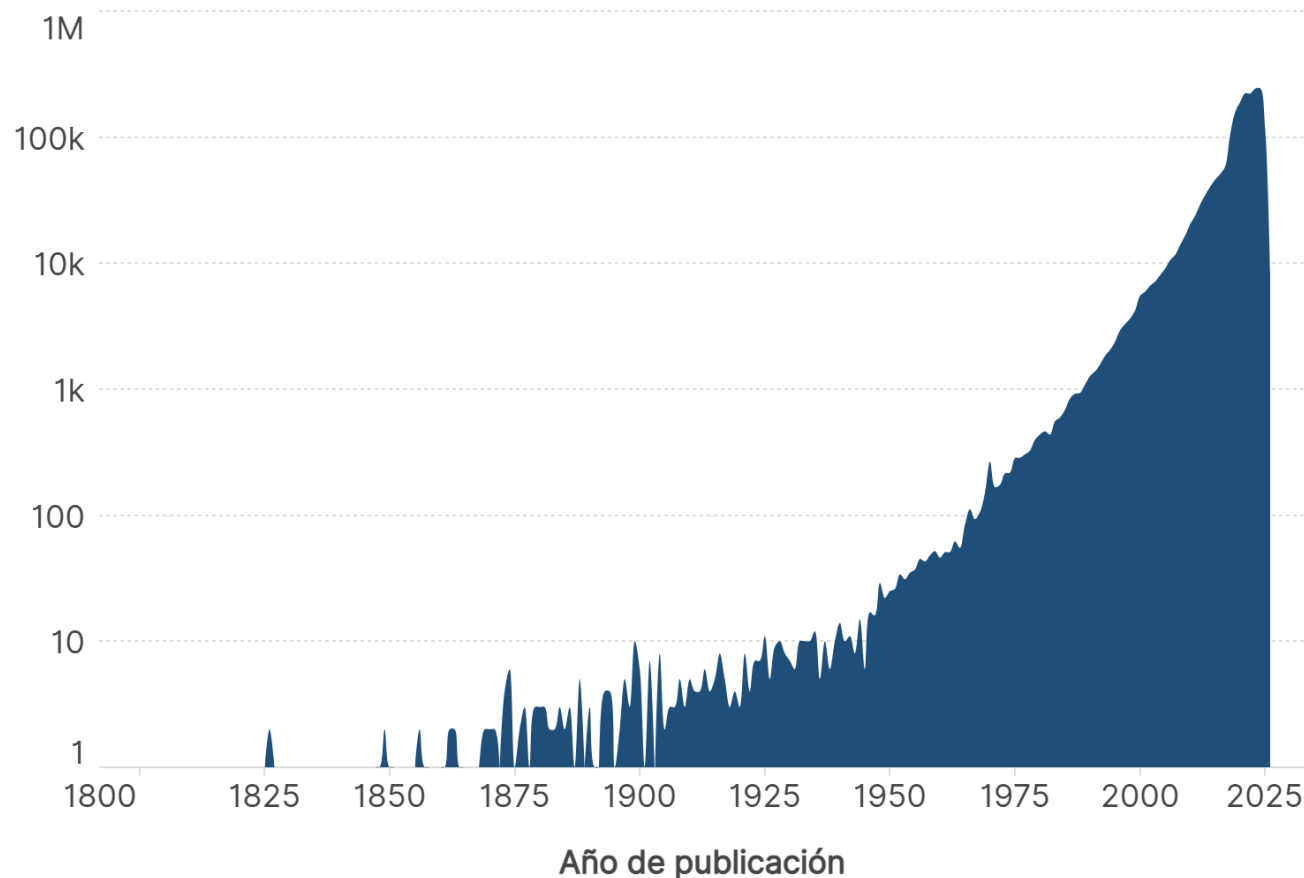
La visibilidad se concentra en una **pequeña fracción de las fuentes**
El **32,9%** de las fuentes aparecen con **una sola publicación citada**.



Publicaciones por año

- La distribución global muestra una fuerte concentración de **publicaciones recientes**.
- Ai2 Asta ha citado 2.357.300 publicaciones, de las cuales el **56,3 % fueron publicadas desde 2020**.
- Este comportamiento puede mejorar la actualización, pero también introduce **sesgos de cobertura y visibilidad**.
- Los patrones globales de la plataforma condicionan los resultados obtenidos en **consultas concretas**.

Número de publicaciones con fuente asignada





5. Ai2 Asta: síntesis de literatura y sistema de citas

Fuentes más visibles

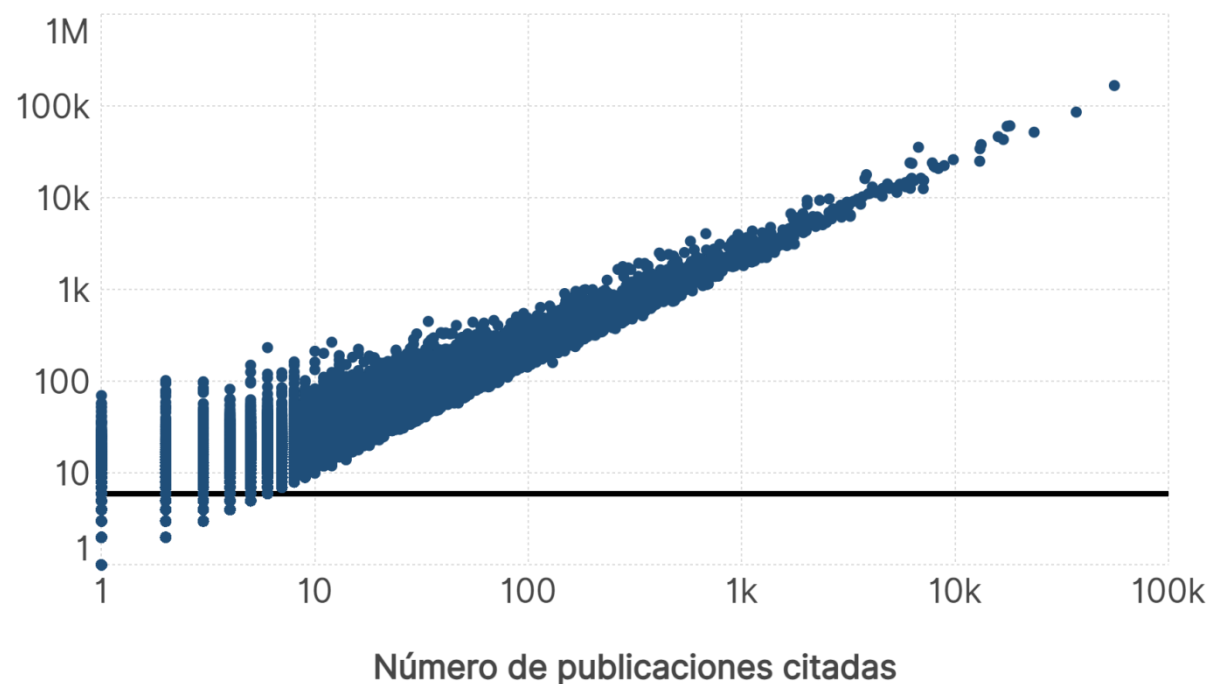
Fuente de publicación	Publicaciones citadas	Apariciones en consultas	Citas
arXiv.org	55.730	124.102	166.921
PLoS ONE	36.936	60.453	85.812
Sustainability	18.050	42.502	60.829
Frontiers in Psychology	17.520	43.295	60.016
Scientific Reports	23.444	36.607	51.798
Int. J. Environmental Research and Public Health	15.914	32.117	46.328
International Journal of Molecular Sciences	16.828	29.510	43.123
IEEE Access	13.241	26.692	37.968
Neural Information Processing Systems	6.732	25.095	35.517
Italian National Conference on Sensors	13.073	23.448	34.234

La correlación de rangos entre apariciones en consultas y citas es muy alta (Spearman $\rho = 0,98$)



Presencia e intensidad de la visibilidad de las fuentes

Número de citas recibidas



Número de citas por publicación



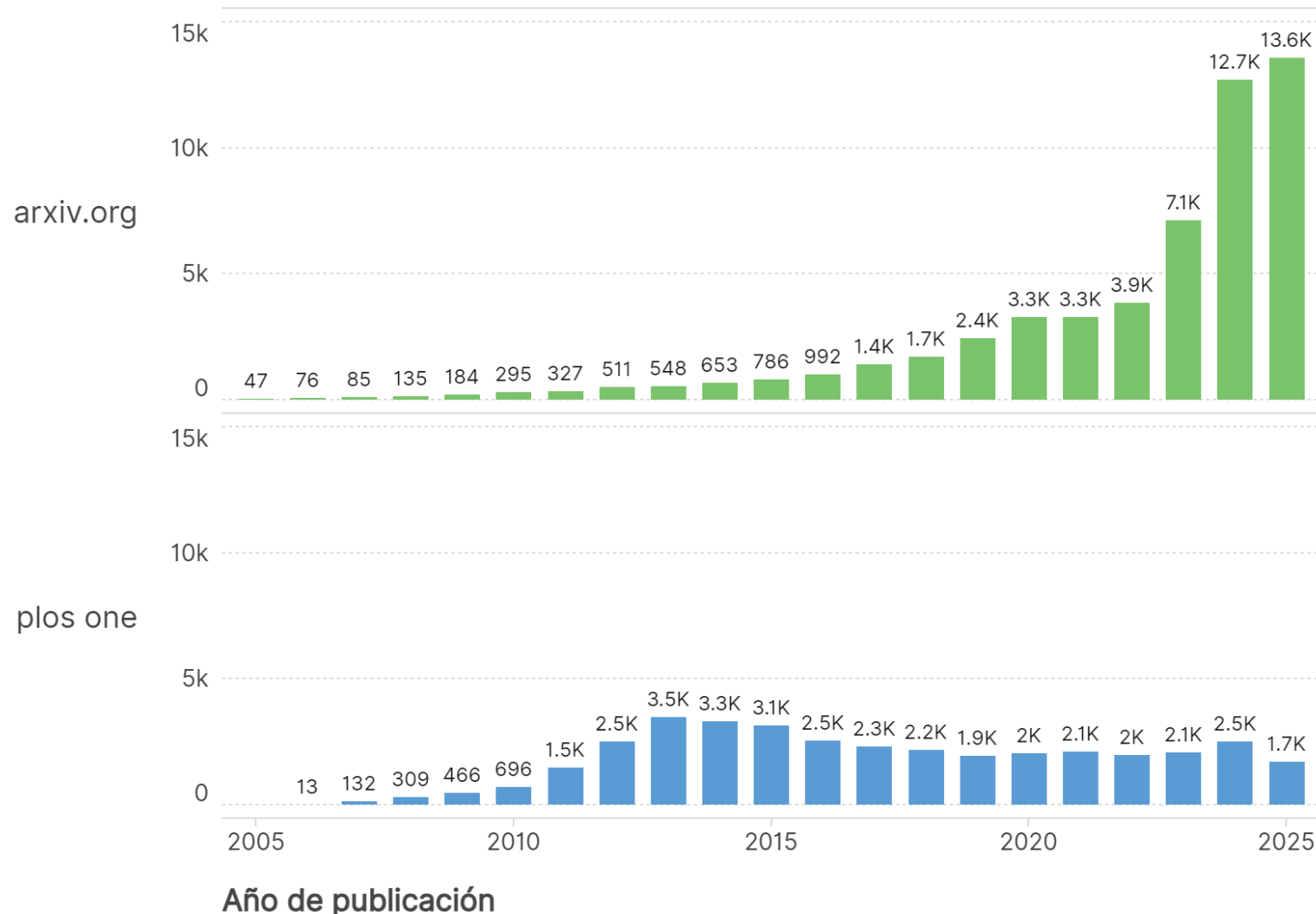
El volumen de citas está fuertemente asociado al número de publicaciones visibles, pero **existen diferencias importantes en la intensidad de citación por publicación**



Trayectorias

- arXiv.org muestra una **trayectoria claramente ascendente** entre 2020 y 2024, tanto en publicaciones como en apariciones en consultas y citas.
- *PLoS One* mantiene una presencia elevada, pero con una **evolución mucho más estable**.

Número de publicaciones citadas por fuente





5. Ai2 Asta: síntesis de literatura y sistema de citas

Trayectorias

Revista	Publicaciones 2020	Publicaciones 2024	Cuota 2020	Cuota 2024
Heliyon	452	3.400	0,24 %	1,39 %
Cureus	462	2.875	0,24 %	1,18 %
Scientific Reports	2.102	3.384	1,11 %	1,39 %
Frontiers in Public Health	172	791	0,09 %	0,32 %
Journal of Clinical Medicine	543	1.191	0,29 %	0,49 %

Varias revistas muestran aumentos claros en presencia relativa. **Heliyon** y **Cureus** destacan por el fuerte crecimiento absoluto y por el aumento de su cuota entre 2020 y 2024



5. Ai2 Asta: síntesis de literatura y sistema de citas

Presencia local

Revista ecuatoriana	Artículos	Apariciones en consultas	Citas
Maskana	7	12	18
Ciencia UNEMI	6	8	14
Chakiñan, Revista de Ciencias Sociales y Humanidades	5	8	13
Ingenius	4	8	14
Rev. de Investigaciones en Energía, Medio Ambiente y Tecnología	4	5	7
INNOVA Research Journal	4	4	5
Iuris Dictio	4	4	5
Alteridad	3	5	6
ACI Avances en Ciencias e Ingenierías	3	3	4
Revista San Gregorio	3	3	3
Eca Sinergia	2	4	6

Se muestran algunos ejemplos identificados al contrastar las fuentes de *Asta* con las 107 revistas de Ecuador incluidas en DOAJ. La identificación no siempre es directa debido a las variantes utilizadas por *Semantic Scholar*.



- Ai2 Asta actúa como un **intermediario** fuerte entre la pregunta del investigador y la literatura visible.
- Sus informes son **estructurados y apoyados en citas**, pero las referencias muestran **inestabilidad**.
- La selección final de **documentos no es plenamente transparente ni reproducible**.
- Estos sistemas pueden influir en qué **publicaciones, revistas y fuentes** ganan visibilidad (*retrieval lottery*).
- El reto no es solo usar estas herramientas, sino comprender **cómo seleccionan, citan y legitiman evidencia**.



- **Aplicar estudios longitudinales** para conocer cómo cambian los resultados tras actualizaciones del sistema.
- **Auditar** el paso de **documentos recuperados** a documentos finalmente citados.
- **Evaluar sesgos de cobertura**: acceso abierto, idioma, campo, revista y repositorio.
- **Desarrollar** criterios de **uso responsable** para investigación, publicación y evaluación científica.



Acción	Alcance	Tarea
Medir	<u>Monitorizar</u> la visibilidad en asistentes de investigación	Comprobar periódicamente si los artículos de la revista son recuperados y citados , para qué consultas, con qué intensidad y cómo evoluciona su presencia.
Facilitar	<u>Cuidar</u> la infraestructura de descubrimiento	Revisar consistencia del título de la revista, DOI/ISSN, metadatos, resúmenes, referencias, interoperabilidad y disponibilidad en las infraestructuras que alimentan estos sistemas.
Auditar	<u>No asumir</u> que indexación es igual a visibilidad	Comparar qué documentos podrían ser recuperados con cuáles terminan siendo seleccionados y citados. Vigilar sesgos por antigüedad, idioma, acceso, temática o fuente.
Interpretar	<u>No confundir</u> visibilidad algorítmica con calidad	Una mayor presencia en respuestas generadas no constituye por sí misma evidencia de mayor calidad o impacto.

Orduña-Malea, E., & Lopezosa, C. (2025). Unraveling the Ai2 Asta Scholarly Research Assistant Citation System.

Revista Panamericana de Comunicación, 7(2), 3675.

<https://doi.org/10.21555/rpc.v7i2.3675>

Tay, A. (2025a). Why I Think Academic Deep Research—or at Least Deep Search—Will “Win”.

<https://aarontay.substack.com/p/why-i-think-academic-deep-research>

Tay, A. (2025b). How Agentic Are Academic Deep Research Tools?

<https://aarontay.substack.com/p/how-agentic-are-academic-deep-research>



¡Muchas gracias por la atención!